Chương 2: Đại lượng ngẫu nhiên

Ví dụ

- Tung cùng lúc 2 xúc xắc. Gọi X là tổng số nút mỗi lần tung. X nhận giá trị ngẫu nhiên trong tập hợp $\{2, 3, ..., 12\}$.
- Cài đặt 1 phần mềm vào máy vi tính. Gọi X là thời gian cài đặt phần mềm đó vào 1 máy tính bất kì. X nhận giá trị ngẫu nhiên.

Khái niệm

- Xét phép thử τ với không gian mẫu Ω.
- Đại lượng ngẫu nhiên $X : \Omega \to \mathbb{R}$, là một *hàm số*, gán mỗi phần tử trong không gian mẫu với một số thực x_i.
- Đại lượng ngẫu nhiên thường được ký hiệu bằng chữ cái in hoa: $X, Y, Z, \ldots$
- Đại lượng ngẫu nhiên *(Random Variable)* còn được gọi là *biến ngẫu nhiên*, hoặc *hàm ngẫu nhiên*
- Ký hiệu $\{X = x\}$ là một biến cố bao gồm *các biến cố sơ cấp* w sao cho $X(w) = x$, nghĩa là

$$\{X = x\} = \{w \in \Omega \colon X(w) = x\}$$

Tương tự, ta có

$$\{X \leq x\} = \{w \in \Omega \colon X(w) \leq x\}$$

$$\{X < x\} = \{w \in \Omega \colon X(w) < x\}$$

Ví dụ

- Tung một đồng xu 2 lần, ta có không gian mẫu:

$$\Omega = \{NN, NS, SN, SS\}$$

- Nếu gọi X là ĐLNN thể hiện *số lần mặt ngửa xuất hiện*, thì:

$$X(NN) = 2,\ X(NS) = X(SN) = 1, X(SS) = 0$$

- Miền giá trị của ĐLNN X: $X(\Omega) = \{0, 1, 2\}$
- Xác suất
  $\{X = 0\} = \{SS\} \Rightarrow P(X = 0) = 1/4,$
  $\{X = 1\} = \{NS, SN\} \Rightarrow P(X = 1) = 1/2,$
  $\{X \leq 1\} = \{NS, SN, SS\} \Rightarrow P(X \leq 1) = 3/4.$

Hàm phân phối xác suất

Định nghĩa: Hàm phân phối xác suất của một ĐLNN X là hàm số

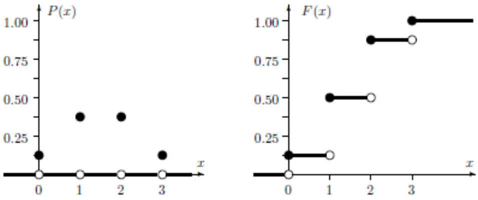
$$F(x) = P(X \leq x) \quad \forall\, x \in \mathbb{R}$$

Tính chất:

- $0 \leq F(x) \leq 1$
- $F(x_1) \leq F(x_2)$ if $x_1 < x_2$
- $\lim\limits_{x \to +\infty} F(x) = F(+\infty) = 1$
- $\lim\limits_{x \to -\infty} F(x) = F(-\infty) = 0$

ĐLNN rời rạc

- ĐLNN rời rạc: X chỉ nhận một số hữu hạn các giá trị hoặc một số vô hạn đếm được các giá trị.
- *Ví dụ 1:* Tung 3 đồng xu,gọi X là số mặt ngửa xuất hiện, thì X là ĐLNN rời rạc



- ĐLNN rời rạc là ĐLNN có đồ thị hàm phân phối dạng bậc thang.

ĐLNN liên tục

- ĐLNN liên tục: *miền giá trị* của X có chứa một khoảng số thực (a,b) (do đó là vô hạn và không đếm được).
- *Ví dụ:* Chọn ngẫu nhiên một số thực trong khoảng [0,1], gọi X là giá trị được chọn. Khi đó X là ĐLNN liên tục.

  VD ĐLNN liên tục: thời gian cài đặt phần mềm, thời gian download file, ...
  VD ĐLNN liên tục: cân nặng, chiều cao, nhiệt độ,...
- Chú ý:
  - Hàm phân phối xác suất của ĐLNN liên tục là hàm liên tục và đạo hàm của nó tồn tại và liên tục hầu khắp nơi.
  - $\boxed{P(x) = P(X = x) = 0 \quad \forall\ x}$

Hàm phân phối và Hàm khối xác suất của ĐLNN rời rạc

Xét ĐLNN *rời rạc* X trên không gian mẫu Ω.

$$X(\Omega) = \{x_1, x_2, x_3, \ldots\} \qquad (x_i < x_{i+1})$$

Hàm phân phối *(Cumulative Distribution Function – CDF)*

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(x_i)$$

Hàm khối xác suất *(Probability Mass Function – PMF)*:

$$P(x) = P(X = x)$$

Tính chất:

1. $P(x) \geq 0, \ \forall x \in X(\Omega)$
2. $P(x) = 0$ nếu $x \notin X(\Omega)$
3. $P(X = x_i) = P(X \leq x_i) - P(X \leq x_{i-1}) = F(x_i) - F(x_{i-1})$
4. $\sum_{\forall i} P(x_i) = 1$

Biểu diễn phân phối xác suất

Phân phối xác suất của một ĐLNN rời rạc có thể được biểu diễn qua

1. Bảng phân phối xác suất:
2. Hàm khối xác suất $P(x)$ (PMF)
3. Đồ thị phân bố *(Histogram)*

Ví dụ

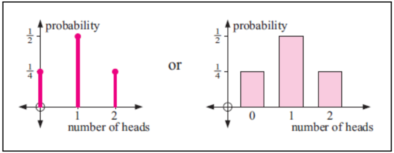
X: số lần xuất hiện mặt ngửa trong phép thử tung một đồng xu 2 lần.

1. Bảng phân phối xác suất:

x	0	1	2
$P(x) = P(X = x)$	$1/4$	$1/2$	$1/4$

2. Hàm khối xác suất: $P(x) = \begin{cases} 1/4 & x = 0 \\ 1/2 & x = 1 \\ 1/4 & x = 2 \\ 0 & \text{otherwise} \end{cases}$

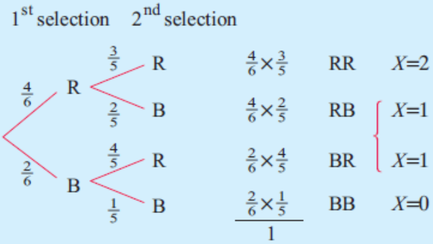
3. Đồ thị



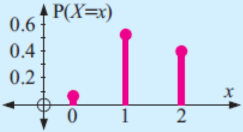
Ví dụ

Example 4

A bag contains 4 red and 2 blue marbles. Two marbles are randomly selected without replacement. If X denotes the number of reds selected, find the probability distribution of X.



x	0	1	2
$P(X = x)$	$\frac{2}{30}$	$\frac{16}{30}$	$\frac{12}{30}$



Hàm khối xác suất PMF

$P(x)$ là hàm khối xác suất của một ĐLNN rời rạc

$$\Rightarrow \begin{cases} P(x) \geq 0 \quad \forall x \\ \sum\limits_{\forall x} P(x) = 1 \end{cases}$$

Ví dụ. Chứng minh các hàm số sau là hàm khối xác suất của một ĐLNN rời rạc

1. $P(x) = \begin{cases} (x^2+1)/34 & \text{nếu } x = 1,2,3,4 \\ 0 & \text{nếu } x \neq 1,2,3,4 \end{cases}$
2. $P(x) = C_3^x (0.6)^x (0.4)^{3-x}, \quad x = 0,1,2,3$

Bài tập

Cho ĐLNN rời rạc X có hàm *khối xác suất* $P(x)$ được xác định bởi công thức

$$P(x) = k\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{4-x} \quad \text{với } x = 0, 1, 2, 3, 4$$

1. Tìm k, từ đó tính $P(X = 2)$;
2. Lập bảng phân phối xác suất của X;
3. Tính $F(3)$ và $P(X > 3)$.

Hàm mật độ xác suất (PDF) của ĐLNN liên tục

Cho X là ĐLNN liên tục và $F(x)$ là hàm phân phối xác suất của ĐLNN X. Khi đó

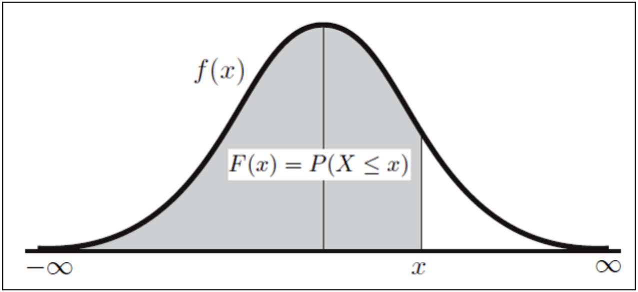
$$\boxed{f(x) = \frac{dF(x)}{dx}}$$

được gọi là hàm mật độ xác suất(Probability density function) của ĐLNN X.

Tính chất:

1. $f(x) \geq 0 \quad \forall\, x \in \mathbb{R}$ và liên tục (hầu khắp nơi).
2. $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a < X \leq b) = \displaystyle\int_a^b f(x)dx = F(b) - F(a)$
4. $F(a) = P(X \leq a) = \displaystyle\int_{-\infty}^{a} f(x)dx$

Hàm mật độ $f(x)$ của ĐLNN liên tục



- Hàm mật độ $f(x) \geq 0$ luôn nằm trên trục hoành.
- Diện tích bên dưới *đường cong mật độ* luôn bằng 1.
- $F(x) = P(X \leq x) =$ diện tích bên trái điểm x.
- $F(x)$ là một hàm *không giảm*.

So sánh với ĐLNN rời rạc và ĐLNN liên tục

ĐLNN rời rạc

- Hàm khối xác suất $P(x)$
  1. $P(x) = P(X = x)$
  2. $P(x) \geq 0 \quad \forall x \in \mathbb{R}$
  3. $\sum P(x) = 1$
- Hàm phân phối $F(x)$:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(x_i)$$

$$\neq P(X < x) \text{ tại vài giá trị } x$$

ĐLNN liên tục

- Hàm mật độ $f(x)$:
  1. $f(x) = F'(x) \geq 0 \quad \forall x$
  2. $P(X = x) = 0 \quad \forall x$
  3. $\int_{-\infty}^{\infty} f(x)dx = 1$
- Hàm phân phối $F(x)$:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

$$= P(X < x) \quad \forall x$$

Chú ý: Một số tài liệu XSTK của Việt Nam gọi hàm khối xác suất là hàm mật độ trong trường hợp ĐLNN rời rạc.

Ví dụ

Cho X là ĐLNN liên tục với hàm mật độ

$$f(x) = \begin{cases} kx & x \in [0,1] \\ 0 & x \notin [0,1] \end{cases}$$

1. Xác định giá trị k.
2. Tìm hàm phân phối $F(x)$.
3. Tính các xác suất $P\left(X < \frac{1}{2}\right)$ và $P\left(\frac{1}{2} < X < \frac{3}{4}\right)$.

Kỳ vọng - Phương sai

Xét ĐLNN X có $X(\Omega)$

- Kỳ vọng: là giá trị trung bình của ĐLNN

$$E(X) = \mu = \begin{cases} \sum_i x_i P(x_i) & \text{nếu } X \text{ rời rạc} \\ \int_{-\infty}^{\infty} x\, f(x) dx & \text{nếu } X \text{ liên tục} \end{cases}$$

- Phương sai:

$$Var(X) = \begin{cases} \sum (x_i - \mu)^2 P(x_i) & \text{nếu } X \text{ rời rạc} \\ \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx & \text{nếu } X \text{ liên tục} \end{cases}$$

Phương sai và độ lệch chuẩn

- Độ lệch chuẩn: $\boxed{\sigma = \sqrt{Var(X)}}$
- Độ lệch chuẩn thường dùng để đo **độ phân tán** của ĐLNN quanh giá trị trung bình $E_X = \mu$
    - Độ lệch chuẩn càng lớn, các giá trị của X càng phân tán xa giá trị trung bình μ;
    - Độ lệch chuẩn càng nhỏ, các giá trị của X càng tập trung gần giá trị trung bình μ.
- Kỳ vọng μ và độ lệch chuẩn σ có *cùng đơn vị* với ĐLNN.

Ví dụ

Cho một kiện hàng có 6 sản phẩm loại 1, 4 sản phẩm loại 2. Chọn ngẫu nhiên có hoàn lại 2 sản phẩm từ kiện hàng này. Gọi X là số sản phẩm loại 1 trong 2 sản phẩm được lấy ra.

1. Tìm luật phân phối xác suất của X
2. Tìm kỳ vọng $E(X)$
3. Tìm phương sai $Var(X)$

Giải ví dụ

1) Gọi A_i là biến cố lần thứ i chọn ngẫu nhiên được một sản phẩm loại 1 từ kiện hàng.

Vì chọn có hoàn lại nên: $\begin{cases} P(A_i) & = \frac{6}{10} = 0,6 = p \\ P(\overline{A_i}) & = 1 - p = 0,4 = q \end{cases}$

X là số sản phẩm loại 1 có trong hai sản phẩm lấy ra

Luật phân phối xác suất của X:

x	0	1	2
$P(x)$	0,16	0,48	0,36

- Vì chọn có hoàn lại nên A_1, A_2 là hai biến cố độc lập:
  $P(X = 0) = P(\overline{A_1}.\overline{A_2}) = P(\overline{A_1}).P(\overline{A_2}) = q^2 = 0,16$
- $P(X = 1) = pq + qp = 0,48$
- $P(X = 2) = p^2 = 0,36$

Giải ví dụ

2) Kỳ vọng (trung bình):

$$\begin{aligned} E(X) = \mu = \sum xP(x) \\ = 0,16 \times 0 + 0,48 \times 1 + 0,36 \times 2 = 1,2 \end{aligned}$$

Vậy, trong 2 sản phẩm chọn ra, *trung bình* có 1.2 sản phẩm loại 1.

3) Phương sai:

$$\begin{aligned} Var(X) &= \sum (x - \mu)^2 P(x) \\ &= 0,16.(0 - 1,2)^2 + 0,48.(1 - 1,2)^2 + 0,36.(2 - 1,2)^2 \\ &= 0,48 \end{aligned}$$

Bài tập

Một doanh nghiệp đầu tư phát triển một loại sản phẩm mới, xác suất thành công là 30%. Chi phí đầu tư bỏ ra là 100 nghìn USD. Nếu không thành công thì mất chi phí đầu tư mà không thu về được gì, nhưng nếu thành công thì thu về được 1 triệu USD (trước khi trừ đi chi phí đầu tư). Tính kỳ vọng lợi nhuận từ dự án đầu tư này.

Tính chất kỳ vọng

Theorem 1

1. $E(a) = a$ *với a là hằng số bất kỳ*
2. $E(aX) = aE(X)$ *với a là hằng số bất kỳ*
3. $E\left(f(X) + g(X)\right) = E\left(f(X)\right) + E\left(g(X)\right)$ *với f, g là hàm số bất kỳ.*
4. $E(aX + bY) = aE(X) + bE(Y)$.

Ví dụ:

1. $E(5) = 5$
2. $E(3X) = 3E(X)$
3. $E(X^2 + 3X + 5) = E(X^2) + 3E(X) + 5$
4. $E(3X + 4Y) = 3E(X) + 4E(Y)$

Tính chất phương sai

1. $Var(X) = E(X^2) - (E(X))^2$
2. $Var(a) = 0$ với a là hằng số bất kỳ
3. $Var(aX) = a^2 Var(X)$ với a là hằng số bất kỳ
4. $Var(aX + b) = a^2 Var(X)$ với a, b là hằng số bất kỳ
5. $Var(X \pm Y) = Var(X) + Var(Y)$ nếu X, Y độc lập

Ví dụ: Cho X là ĐLNN có trung bình 8.1 và độ lệch chuẩn 2.37. Tính trung bình và độ lệch chuẩn của $Y = 4X - 7$.

Ví dụ

Một ĐLNN X chỉ nhận 1 trong 2 giá trị $-a$ và a $(a > 0)$, mỗi giá trị với xác suất 50%. Tính giá trị kỳ vọng, phương sai và độ lệch chuẩn của X.

- $E(X) = \mu = \sum xP(x) = (-a).50\% + a.50\% = 0$
- $Var(X) = \sum (x-\mu)^2 P(x) = (-a-0)^2.50\% + (a-0)^2.50\% = a^2$

  hoặc:

$$\begin{aligned} Var(X) = E(X^2) - (E(X))^2 &= \sum x^2 P(x) - \mu^2 \\ &= \left[(-a)^2.50\% + a^2.50\%\right] - 0^2 = a^2 \end{aligned}$$

- Độ lệch chuẩn: $\sigma = \sqrt{Var(X)} = a$.

Mode - Giá trị tin chắc nhất

- Định nghĩa: Mode của một ĐLNN *rời rạc* là giá trị của ĐLNN mà tại đó *xác suất xảy ra là lớn nhất*.

$$Mod(X) = x^* \Leftrightarrow P(X = x^*) \text{ max}$$

  Mode của một ĐLNN *liên tục* là giá trị của ĐLNN mà tại đó hàm mật độ $f(x)$ đạt giá trị cực đại.

- Ví dụ: Một hộp đựng 4 quả cầu giống nhau đánh số 1, 2, 3, 4. Lấy ngẫu nhiên 2 quả. Gọi Y là tổng của 2 số ghi trên hai quả đó, ta có bảng phân phối xác suất của Y:

Y	3	4	5	6	7
P_Y	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- $Mod(Y) = 5$ vì $P(Y = 5)$ là *lớn nhất*.

Trung vị (Median)

- Định nghĩa: Trung vị của ĐLNN *rời rạc* X, ký hiệu Med(X), là giá trị x_0 sao cho:

$$P(X < x_0) \leq \frac{1}{2} \quad \text{và} \quad P(X > x_0) \leq \frac{1}{2}$$

  Trung vị của ĐLNN *liên tục* là giá trị x_0 của ĐLNN mà tại đó $F(x_0) = 1/2$, nghĩa là *đường cong mật độ* bị tách thành 2 phần có diện tích bằng nhau

$$Med(X) = x_0 \Leftrightarrow P(X < x_0) = P(X > x_0) = \frac{1}{2}$$

- **Lưu ý.** để tìm $Med(X) = x_0$ của ĐLNN liên tục, ta giải pt

$$F(x_0) = 1/2 \qquad \text{hoặc} \qquad \int_{-\infty}^{x_0} f(x)\,dx = 1/2$$

Ví dụ

Ví dụ 1: Giả sử ĐLNN Y có phân phối xác suất như sau

Y	3	4	5	6	7
P_Y	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

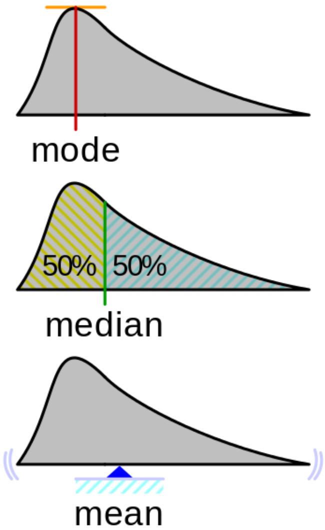
$$\text{Med}(Y) = 5 \text{ vì } P(Y < 5) = P(Y > 5) = \frac{2}{6} < \frac{1}{2}.$$

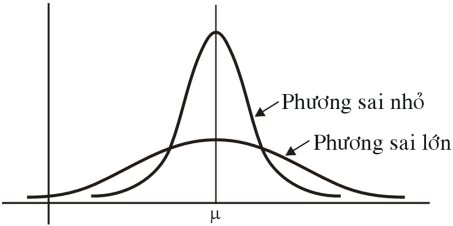
Ví dụ 2:Cho ĐLNN X có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Tìm các giá trị

1. Kỳ vọng $E(X)$
2. Phương sai $Var(X)$ và độ lệch chuẩn σ_X
3. $Mod(X)$ và $Med(X)$


mode
50%
50%
median
mean


Phương sai nhỏ
Phương sai lớn
μ

Phân phối nhị thức (Binomial distribution)

Định nghĩa 1: Xét dãy n phép thử Bernoulli với xác suất xảy ra biến cố A trong mỗi lần thử là p. Gọi X là số lần biến cố A xảy ra trong n phép thử đó. Khi đó X là ĐLNN rời rạc có

$$X(\Omega) = \{0, 1, 2, \ldots, n\}$$

$$P(X = k) = C_n^k p^k (1-p)^{n-k}, \qquad x = 0, 1, 2, \ldots, n$$

Ta nói X có phân phối nhị thức và kí hiệu $X \sim B(n, p)$.

Định nghĩa 2: Xét tổng thể (kích thước rất lớn) mà tỷ lệ phần tử có tính chất A là p. Chọn ngẫu nhiên n phần tử. Gọi X là số phần tử có tính chất A trong n phần tử. Khi đó X có phân phối Nhị thức.

$$X \sim B(n, p)$$

$$P(X = k) = C_n^k p^k (1-p)^{n-k}, \qquad x = 0, 1, 2, \ldots, n$$

Phân phối nhị thức

Gọi X là số lần biến cố A xảy ra trong n phép thử đó.
$X \sim B(n, p)$

$$\begin{aligned}
n &= \text{số lần thử} \\
p &= \text{xác suất biến cố A xảy ra trong mỗi lần thử} \\
P(x) &= C_n^x p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n \\
F(x) &= \begin{cases} 0 & \text{if } x < 0 \\ \sum\limits_{0 \le k \le x} C_n^k p^k (1-p)^{n-k} & \text{if } 0 \le x \le n \\ 1 & \text{if } x > n \end{cases} \\
E(X) &= np \\
Var(X) &= np(1-p)
\end{aligned}$$

Ví dụ

1. Một kĩ thuật viên theo dõi 12 máy tự động. Xác suất để mỗi máy cần đến sự điều chỉnh của kĩ thuật viên trong mỗi ngày là 0.1. Các máy này hoạt động độc lập nhau.
   a) Gọi X là số máy cần sự điều chỉnh mỗi ngày. Tìm phân phối xác suất của X.
   b) Tìm xác suất mỗi ngày có ít nhất 1 máy cần điều chỉnh.
2. Xác suất trúng mỗi tờ vé số là 0.001.
   a) Hỏi nếu mua 20 tờ vé số thì xác suất trúng 2 tờ là bao nhiêu ?
   b) Hỏi trung bình mua 20 tờ vé số thì trúng bao nhiêu tờ?

1. Có một trò chơi điện tử được phát hành. 60% người chơi hoàn thành được hết các level và có 30% trong số đó sau đó mua phiên bản tiếp theo. Hỏi trung bình trong 15 người chơi, có bao nhiêu người mua phiên bản tiếp theo ? Xác suất có ít nhất 2 người chơi sẽ mua phiên bản mới trong 15 người chơi ?

Phân phối hình học

Xác suất trúng khi bắn một viên đạn là p. Ta bắn vào mục tiêu trong điều kiện như nhau cho đến khi có 1 viên trúng thì dừng. Gọi X là số đạn cần bắn. Tìm phân phối của X.

Xét dãy phép thử Bernoulli với xác suất biến cố A xảy ra trong mỗi lần là p. Gọi X là số lần cần thử để biến cố A xảy ra đầu tiên. Ta nói X là ĐLNN có phân phối hình học.

$$
\begin{aligned}
p &= \text{xác suất thành công mỗi lần thử} \\
P(x) &= (1-p)^{x-1}p, \quad x = 0, 1, 2, \ldots \\
E(X) &= \frac{1}{p} \\
Var(X) &= \frac{1-p}{p^2}
\end{aligned}
$$

Định nghĩa Phân phối siêu bội

Xét tập hợp gồm N phần tử trong đó có M phần tử có tính chất $\wp$. Chọn *không hoàn lại* ngẫu nhiên n phần tử và gọi X là số phần tử có tính chất $\wp$ trong n phần tử chọn ra, thì X có phân phối siêu bội, ký hiệu $X \sim H(N, M, n)$.

- Miền giá trị: $X(\Omega) = \{0, 1, ..., \min(n, M)\}$
- Hàm khối xác suất: $P(X = x) = \dfrac{C_M^x C_{N-M}^{n-x}}{C_N^n}$

Chứng minh.

- Chọn n phần tử bất kỳ từ N phần tử: C_N^n.
- Chọn x phần tử có tính chất $\wp$ (từ M phần tử): C_M^x;
  Chọn $(n-x)$ phần tử không có tính chất $\wp$ (từ $(N-M)$ phần tử): C_{N-M}^{n-x}.

Theo định nghĩa xác suất cổ điển, ta có:

$$P(X=x)=\frac{C_M^x C_{N-M}^{n-x}}{C_N^n}$$

$\square$

Tham số đặc trưng của $X \sim H(N, M, n)$

- Kỳ vọng (giá trị trung bình):

$$\boxed{E(X) = \mu = np}$$

  với $p = M/N$.

- Phương sai:

$$\boxed{Var(X) = np(1-p)\frac{N-n}{N-1}}$$

- Độ lệch chuẩn:

$$\boxed{\sigma = \sqrt{Var(X)}}$$

Một máy tính có 20 thư mục trong đó có 5 thư mục bị nhiễm virus. Chọn ngẫu nhiên 6 thư mục để kiểm tra xem có bị nhiễm virus hay không.
a) Tính xác suất có ít nhất 3 thư mục bị nhiễm virus.
b) Trung bình có bao nhiêu thư mục bị nhiễm trong 6 thư mục được chọn ?

Định nghĩa Phân phối Poisson

Một ĐLNN X được gọi là có **phân phối Poisson** với tham số λ, nếu các giá trị của nó là các số nguyên không âm, và có hàm khối xác suất

$$P(x) = P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda} \qquad \text{với } x = 0, 1, 2, \ldots$$

Ký hiệu: $X \sim P(\lambda)$.

- Miền giá trị: $X(\Omega) = \{0, 1, 2, \ldots\}$
- Kỳ vọng: $E(X) = \lambda$
- Phương sai: $Var(X) = \lambda$

ĐLNN có phân phối Poisson

Mô hình phân bố Poisson thường được dùng cho các ĐLNN có dạng *"số sự kiện xảy ra trong một khoảng thời gian, hoặc một khoảng không gian nào đó"*.

- Số cuộc gọi điện đến một trạm điện thoại trong 1 giờ.
- Số lỗi in sai trên một trang sách trong một quyển sách.
- Số lượng cá bị đánh bắt trên 1 hồ lớn trong một ngày.
- Số tai nạn giao thông trên một con đường trong một tháng...

Tham số λ chính là giá trị trung bình (kỳ vọng) của ĐLNN.

Ví dụ. Nếu tại một con đường, trung bình một tháng có 7.3 vụ tai nạn giao thông, và X là số vụ tai nạn trong một tháng, thì $X \sim P(\lambda = 7.3)$.

Phân phối Poisson

Ví dụ

Trung bình 5 phút có 15 người vào một đại lý bưu điện. Tính xác suất trong 1 phút có 4 người vào đại lý bưu điện đó. Biết rằng số người vào đại lý là một ĐLNN rời rạc có phân phối Poisson.

- Số người vào trung bình mỗi phút: $15/5 = 3$ người.
- Gọi X là số người vào trạm bưu điện trong 1 phút, thì $X \sim P(\lambda = 3)$.

$$P(4) = P(X = 4) = \frac{3^4 e^{-3}}{4!} \approx 16.8\%$$

Một công ty cung cấp dịch vụ internet có trung bình 20 khách hàng đăng kí tài khoản mới mỗi ngày. Biết rằng số tài khoản mới đăng kí mỗi ngày có phân phối Poisson.
a) Tính xác suất có ít nhất 3 tài khoản mới đăng kí ra trong ngày mai.
b) Tính xác suất có ít nhất 6 tài khoản mới đăng kí trong 2 ngày tới.

Bài tập ôn

1.17 Có 1000 vé số trong đó có 20 vé trúng thưởng. Một người mua 30 vé, tìm xác suất để người đó trúng 5 vé.

1.18 Để được nhập kho, sản phẩm của nhà máy phải qua 3 vòng kiểm tra chất lượng độc lập nhau. Xác suất phát hiện ra phế phẩm ở các vòng lần lượt theo thứ tự là 0,8; 0,9 và 0,99. Tính xác suất phế phẩm được nhập kho.

1.19 Một thủ kho có một chùm chìa khóa gồm 9 chiếc trông giống hệt nhau trong đó chỉ có một chiếc mở được kho. Anh ta thử ngẫu nhiên từng chìa khóa một, chiếc nào được thử thì không thử lại. Tính xác suất anh ta mở được cửa ở lần thử thứ 4.

1.20 Một lô hàng có 9 sản phẩm. Mỗi lần kiểm tra chất lượng lấy ngẫu nhiên 3 sản phẩm. Sau khi kiểm tra xong trả lại vào lô hàng. Tính xác suất để sau 3 lần kiểm tra lô hàng, tất cả các sản phẩm đều được kiểm tra.

1.21 Một nhà máy ôtô có ba phân xưởng I, II, III cùng sản xuất ra một loại pít-tông. Phân xưởng I, II, III sản xuất tương ứng 36%, 34%, 30% sản lượng của nhà máy, với tỷ lệ phế phẩm tương ứng là 0,12; 0,1; 0,08.

a) Tìm tỷ lệ phế phẩm chung của nhà máy.

b) Lấy ngẫu nhiên một sản phẩm kiểm tra và được sản phẩm là phế phẩm. Tính xác suất để phế phẩm đó là do phân xưởng I, II, III sản xuất.